10-1-2016

Wage Determination in Social Occupations: the Role of Individual Social Capital

Julie L. Hotchkiss
Federal Reserve Bank of Atlanta, julie.l.hotchkiss@atl.frb.org

Anil Rupasingha
U.S. Department of Agriculture, anil.rupasingha@ers.usda.gov

W. J. Usery Workplace Research Group Paper Series

Working Paper 2016-10-1
October 2016

Wage Determination in Social Occupations: The Role of Individual Social Capital

Julie Hotchkiss
Federal Reserve Bank of Atlanta
Georgia State University

Anil Rupasingha
U.S. Department of Agriculture

Georgia State University
Andrew Young School of Policy Studies

Wage Determination in Social Occupations: the Role of Individual Social Capital

Julie L. Hotchkiss
Federal Reserve Bank of Atlanta and Georgia State University
Research Department
1000 Peachtree St. NE
Atlanta, GA 30303
404-498-8198
Julie.L.Hotchkiss@atl.frb.org

Anil Rupasingha
U.S. Department of Agriculture
Economic Research Service
1400 Independence Ave SW Mail Stop 1800
Washington, D.C. 20024
202-694-5227
anil.rupasingha@ers.usda.gov

October 26, 2016

Abstract

We make use of predicted social and civic activities (social capital) to account for selection into "social" occupations. Individual selection accounts for more than the total difference in wages observed between social and non-social occupations. The role that individual social capital plays in selecting into these occupations and the importance of selection in explaining wage differences across occupations is similar for both men and women. We make use of restricted 2000 Decennial Census and 2000 Social Capital Community Benchmark Survey. Individual social capital is instrumented by distance weighted surrounding census tract characteristics.

JEL classification:
J31 Wage Differentials
J24 Occupational Choice
C34 Switching Regression Models

Key words: social capital, wage differentials, occupational choice, switching regression, non-public data, factor analysis

---

Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau, the Federal Reserve Bank of Atlanta, the U.S. Department of Agriculture, or the Economic Research Service. All results have been reviewed to ensure that no confidential information is disclosed. The authors appreciate comments from and discussions with Melissa R. Banzhaf, Tim Dunne, Barry Hirsch, Lisa Kahn, Tom Mroz, M. Melinda Pitts, and Matthew Ross. Research assistance from Augustine Denteh, Tom Zichong Qu, and Kalee Burns is also appreciated.

Wage Determination in Social Occupations: the Role of Individual Social Capital

1 Introduction and Background

The purpose of this paper is to assess the role that individual social capital plays in the determination of observed wage differentials between "social" and "nonsocial" occupations. As an extra-market characteristic, or reflection of preferences, social capital is expected to have a significant impact on a worker's occupational choice. More specifically, we conjecture that individuals' social capital in the form of "sociability" and "altruism" may play a role in the determination of their wages, directly through the value employers place on these attributes, and indirectly through self-selection into occupations based on preferences reflecting different levels of social capital. The importance of pre-market characteristics and preferences in labor market outcomes, and particularly, occupational choice, is well-established in the literature (e.g. (Wiswall and Zafar 2016; Speer 2017). This paper places particular emphasis on the mechanism of self-selection in determining wages in social and nonsocial occupations and the importance of social capital in that relationship.

Wages paid in occupations labeled as "caring" or "social" have received attention in the economics literature, with the historical focus being the fact that these occupations are dominated by women (e.g., England, Budig, and Folbre 2002; Kilbourne et al. 1994; Pitts 2003). Because of this high representation of women, wage penalties associated with the occupations are often identified as an important source of wage differentials between men and women. Although the wage penalty may be greater for men in these occupations, as it typically found in the literature, there are many more women than men found in these occupations, making the existence of a penalty particularly salient for women.

A natural question arises as to why women, mostly, would continue to choose these occupations if doing so means they pay a penalty in the labor market. Pitts (2003) finds that rather than women being pushed, or segregated, into these occupations, the skills and attributes women bring to the labor market garner greater reward in occupations dominated by women (occupations often classified as caring or social) than they do in occupations dominated by men -- the choice is an economic one.[1] In addition, there is evidence that social skills most often associated with women are growing in importance and value across the entire labor market. Deming (2015) provides evidence that since 1980, jobs that require relatively intense social skills have enjoyed the bulk of the job growth and that high-paying jobs that are not easily replaced by technology increasingly require social skills. The potential implication of this for gender wage differentials is obvious.[2]

In order to classify occupations as "social" and "nonsocial," we appeal to a list of occupations identified exogenously (not by the authors or statistically) as social by an online career counseling web site, Career Key (https://www.careerkey.org/explore-career-options/social-careers-career-clusters.html). We are interested in classifying occupations as "social" based on what someone is thinking (or is told) is a social occupation, rather than based on what activities people in certain occupations report as their tasks. This produces an ex ante classification versus an ex poste classification of occupations. This list of occupations (found in Appendix A) overlaps with those listed by England, Budig, and Folbre (2002) and Hirsch and Manzella (2015) as caring occupations, such as teacher and social worker, but excludes those that might be considered more entrepreneurial, such as doctors and lawyers. The list also

[1] Of course, why women's attributes aren't valued as highly in occupations dominated by men is, in itself, a valid research question.
[2] Also see (Cortes, Jaimovich, and Siu 2016; Deming and Kahn 2016).

overlaps with those that would be considered by Deming (2015) and Ngom (2003) to rely heavily on social skills, such as physical therapist, but exclude those that arguably don't necessarily contribute to the betterment of others, such as cashier and manager.

In addition, we make use of O*NET to validate that the occupations we classify as "social" as being ones that rank highly in attributes commonly associated with occupations high in sociability and/or caring. As expected, occupations identified as social for the purposes of this paper are high in attributes such as communicating with others, providing service and working in a cooperative environment, and low in attributes such as compensation, working alone, and using machinery (see Appendix A).

We take a gender-neutral approach by simply performing separate analyses for men and women. Separate analyses by gender makes sense as the literature generally finds that wage determination in a similar class of caring or social occupations differs by gender (for example, see Hirsch and Manzella (2015) and Kilbourne et al. (1994). It will also allow us to determine whether selection into social and non-social occupations differs by gender and to identify any differences in the role individual social capital might play in those choices. Individual social capital can be thought of as part of a more general class of pre-market characteristics (as in Gould 2002; Speer 2017; Heckman, Stixrud, and Urzua 2006; Kuhn and Weinberger 2005) that are expected to impact an individual's occupational choice.

We explore the importance of two dimensions of social capital: altruism and sociability. As part of his measure of pre-market social skills (Speer 2017) identifies a respondents' participation in clubs and sports in high school as indicators of high sociability; we are able to account for current altruistic and social activity, rather than proxy for it with historical behavior. The analysis in this paper will allow us to identify the importance of individual selection (as

determined by individual social capital) into social and non-social occupations; assess how that selection figures into wage determination in those occupations, independently of any direct influence on wages social capital might have; and to identify any differences for men and women.

1.1 Social Capital

Numerous studies across various disciplines have established that social capital, or civic engagement, plays a positive role in economic, community, and social development in a society. The concept of social capital gained eminence in social science with the much-publicized work of Putnam (1995a, 1995b, 1994); see also Coleman (1988) and Wollcock (2001). Putnam (1995a, 19) defines social capital as interactions among individuals through social networks that lead to norms of reciprocity and trustworthiness. Coleman (1988, 598) defines social capital as, "a variety of different entities, with two elements in common: they all consist of some aspect of social structure, and they facilitate certain actions of actors – whether personal or corporate actors – within the structure." Wollcock (2001) describes social capital as norms and networks that facilitate collective action in the society. And, perhaps a little more tangibly, and, certainly, more historically, Hanifan (1916, 130), in his lament about its near total absence in rural school districts and how one community successfully developed it, defines social capital as, "that in life which tends to make...tangible substances count for most in the daily lives of people, namely, goodwill, fellowship, mutual sympathy, and social intercourse among a group of individuals and families who make up a social unit."

Most of the studies of labor market outcomes that might fall under the social capital literature umbrella have been more concerned with the operation of social and professional networks and job contacts (who knows who), rather than on social and civic activities and

engagement (for example, see Holzer 1987; Montgomery 1991; Lin 1999; Bayer, Ross, and Topa 2005; Hellerstein, McInerney, and Neumark 2008; Schmutte 2015; Beaman 2012; Bentolila, Michelacci, and Suarez 2010; Mouw 2003; Wegener 1991; Cingano and Rosolia 2012; Hensvik and Skans 2016). Examples of exceptions to this generalization are Saffer and Lamiraud (2008) who investigate the relationship between hours of work and social interaction with friends and family and Ngom (2003) who performs an analysis similar to that in this paper, although limited to the analysis of men only; he finds no relationship between social capital investments (instrumented by the number of club or group activities and the number of hours spent on social activities) and either selection into social occupations or earnings in those occupations.

The analysis of this paper contributes to the social capital literature by identifying a link between social and altruistic behavior and supply of labor to occupations aligned with those behaviors. Social capital may turn out to be an important characteristic heretofore rarely considered in this context.

1.2 Occupational Choice

The primary contribution of the analysis in the paper to the occupational choice literature is the use of individual social capital measures as an important extra-market attribute that may signal an individual's comparative advantage in occupations classified as "social." The analysis in this paper builds on the ideas based in Roy (1951) and formalized by Heckman and Honoré 1990), Heckman and Sedlacek (1990), and Gould (2002), among others, promoting the idea of occupational choice as an exercise in comparative advantage. Mani and Mullin (2004) raise the question of whether social perception may also guide workers' occupational choice, although we will have no way of assessing that incentive in this paper; we will assume that workers seek out those occupations in which their particular package of attributes will reap the greatest return.

Unlike some of the previous literature that focuses on explicit returns to certain individual skills in the labor market (e.g., Kilbourne et al. 1994; Hirsch and Manzella 2015; Speer 2017; Deming 2015), this paper is concerned with how certain attributes might affect wages through their determination of selection into certain occupations, with a focus on occupations differentiated by whether they are considered "social" or "non-social." However, in light of these earlier studies, we will also include measures of social capital directly in the wage equation itself.

By allowing an individual's level of social capital to be a determinant in his/her occupational choice and a determinant of the wage directly, we are able to identify both the value placed on those characteristics by employers as well as the role the characteristics play in an individual's decision to choose an occupation for which those characteristics are best suited. (Glaeser, Laibson, and Sacerdote 2002, F450) provide some suggestive evidence that people in occupations that have, "A lot of contact with other people" (i.e., social occupations) also have a higher level of social capital (greater number of personal social interactions and involvement in formal community organizations).

We make use of two measures of an individual's social capital that are likely to be most closely aligned with attributes of social or caring occupations. We measure social capital as evidenced by observed (predicted) adult social engagement and altruistic activities; other efforts in the literature to account for individual "sociability" have used retrospective information on the degree of participation in school clubs and sports (for example, see Speer 2017; Deming 2015; Ngom 2003). Since measures of individual social and altruistic activity are not available in standard labor market surveys, we employ two-sample two-stage least squares (2S2SLS) to obtain a predicted value for measures of social capital for individuals. This is the only analysis of

occupational choice that we know of that makes use of observed, adult social capital measures to model selection into certain occupations.

There is a fairly rich literature that has investigated wage differentials between occupations designated as social or caring and contrasting occupations. Much of the earliest literature was mostly concerned with the importance of women crowding into social or caring occupations, which on average pay lower wages for a given set of human capital characteristics, in the determination of male/female pay gaps. Kilbourne et al. (1994) find a two and four percent penalty for men and women, respectively, per unit of "nurturing" in an occupation. Overall, they conclude that the devaluation of occupations dominated by women accounts for between eight percent and 17 percent of the observed pay gap between men and women. Directly focused more on wage differentials across occupations, England, Budig, and Folbre (2002) uncover a five percent penalty (for both men and women) for being in a caring occupation; the authors point out that even though the penalty is similar across men and women, women suffer more from this penalty because there are many more women than men in caring occupations. Also specifically focused on caring occupations, Hirsch and Manzella (2015) find a greater penalty for men, although it appears to be a fairly modest penalty for both men and women (note that they construct a "caring" index whose units are difficult to interpret).

Most directly related to the methodology employed in this paper, Pitts (2003) explores the penalty for being employed in a female-dominated occupation (vs. non-female dominated), which is relevant since social and caring occupations are significantly dominated by women. She also finds a greater penalty for men than for women, roughly on the order of a 25 percent penalty for men and a small 0.5 percent penalty for women, overall. Although Pitts employs the same

switching regression framework estimated here, we believe that ours is the first estimation of the switching regression with selection into occupational group.

All of these analyses include in one regression a dummy indicator for whether a worker is observed in a social or caring or female-dominated occupation, or include an index measuring the degree of caring or sociability of the occupation. We allow differential selection and differential wage determination in social and non-social occupations. We will estimate models commensurate with previous literature to put some perspective on our final results.

2 Methodology

The methodology used to discern the importance of social capital in the selection of workers into social occupations and the implications of this selection for determination of social and non-social occupational wages has many steps. The first step involves obtaining parameter estimates with which we can predict individual level social capital for respondents in the 2000 Decennial Census (DC). The DC contains the largest sample possible of individual labor market information, but, unfortunately, does not contain any measures of social capital. We make use of the 2000 DC because this is the year of the survey that we will use to obtain predictors of individual social capital. The commonly applied technique of using one sample to obtain predictors for an outcome to be applied to a second sample is called two-sample two-stage least squares (2S2SLS) (Ridder and Moffitt 2007), made popular by Angrist and Krueger 1992, 1995).[3] The second step will be to estimate the selectivity-corrected wage equations for social

[3] Also see Inoue and Solon (2010) for the distinction between 2S2SLS and two-sample instrumental variables (2SIV). Other applications of 2S2SLS can be found in Dee and Evans 2003), Carroll, Dynan, and Krane (2003), Currie and Yelowitz (2000), Fang, Keane, and Silverman (2008), and Keane and Stavrunova (2014).

and non-social occupations using a standard switching-regression model (à la Willis and Rosen 1979; L.-F. Lee 1978).

2.1 Estimating Individual Social Capital

In 2000 the Roper foundation conducted a national survey, the Social Capital Community Benchmark Survey (SCCBS), to gauge the level of a multitude of dimensions of individual social capital. We use the SCCBS as the source for estimating social capital determinants. Parameter estimates are obtained from the SCCBS data that can be used to predict social capital in the DC. Fortunately, the SCCBS and the DC are fairly harmonious with respect to their measures of demographics. This is fortunate, since we are restricted to those variables that are found in both surveys in order to use the estimated parameters from one data set to predict social capital in the second.

2.1.1 Creating Weights to use in the SCCBS. Since we are predicting out of sample, however, and in spite of the fact that both the DC and the SCCBS are both national surveys, we are also interested in how the two samples compare in their distributions across demographics. In other words, we want to be sure that the parameter estimates obtained from the SCCBS sample are likely to be applicable, at least at the means, to observations in the DC. Both the SCCBS and DC surveys contain individual weights designed to generate a random national sample. Using the survey-provided sample weights, 91 percent of the weighted means of the common variables in the DC and SCCBS were statistically different from one another at least at the 95 percent confidence level.

To estimate the social capital equation on a sample that is more representative of the DC (for which the prediction will be made), we use an inverse probability weighting methodology, akin to the one used in DiNardo, Fortin, and Lemieux (1996), in order to create a counterfactual

distribution of the SCCBS that is much more similar to the DC. This amounts to estimating, in the combined DC and SCCBS samples, the probability of an observation being observed in SCCBS, using as explanatory variables as many demographics and their cross-multiples as is feasible:

$$P(observation\ i\ \in SCCBS|X) = \Lambda(X'b)\ . \quad (1)$$

The parameter estimates from this logit model are then used to construct the inverse probability ratio, $\frac{\Lambda(X'\hat{b})}{1-\Lambda(X'\hat{b})}$, for each observation in the SCCBS. This is the re-weighting function used to modify the individual weight provided in the SCCBS.

With over 20 million observations in the DC, even small practical differences in means will be statistically significant. However, there is significant improvement using the inverse-weighted adjustment to the means in the SCCBS. Using the new weights, the percent of common variables that are statistically different from one another is reduced to 67 percent -- with 84 percent of the re-weighted means of all variables being significantly closer to the DC mean than they were using the survey supplied weights.[4]

2.1.2 Identifying a Person's Unobserved Social Capital. A person's social capital is not a characteristic that is observed. In addition, there are many dimensions to social capital, from sociability, altruism, political engagement, etc. (for example, see Putnam 1995a). The SCCBS contains a multitude of questions designed to elicit, based on observed/reported activity, the level of these different dimensions of social capital. In the context of the analysis in this paper, we focus on two measures that are most relevant for the consideration of selection into social occupations -- sociability and altruism -- as opposed to, for example, political engagement or religiosity. We perform a factor analysis using the responses to a variety of questions in the

[4] More details of the means comparisons are suppressed for confidentiality reasons.

SCCBS in order to identify each person's latent degree of sociability and altruism. Factor analysis is designed to uncover the unobserved factor common to observed responses to a number of survey questions. We uncover a person's "altruism" factor based on questions related to community involvement activities, and a person's "sociability" factor based on questions related to the person's social interactions. Details of the factor analysis are found in Appendix B.

2.1.3 Estimating Social Capital. The altruism and sociability factors identified are continuous individual variables. Since we have little hope of adequately fitting those continuous variables, and since the metric of each measure is uninterpretable anyway, we create low, medium, and high categories of each factor by splitting the distribution of the sample into thirds. Since the categories created are ordered from lower to higher levels of the social capital variable in question, we then estimate the parameters of each social capital equation as an ordered logit.[5] The probability that individual *i*, living in census tract *t*, has social capital level *k* of type *j* (*j*=sociability, altruism) is formally expressed as:

$$Pr[SK_{it}^{j} = k] = Pr[\mu_{k-1} < \alpha_0 + \mathrm{A}'X_i + \mathrm{B}'Y_i + \mathrm{C}'Z_t + u_{it} \leq \mu_k] \,, \qquad (2)$$

where u_{it} is assumed to be logistically distributed and the estimated cutpoints μ=1,2 separate three possible outcomes *k*= low, medium, high for each type of social capital.

Regressors X_i reflect person *i*'s demographics (age, education, race/ethnicity, geography, marital status, citizenship, disability status, employment sector and industry) and will also be included as determinants for both occupational choice and wages. Regressors Y_i are additional determinants of a person's level of social capital (children, household income, length of time in area), and will also be included in the person's occupational choice equation, but excluded from wage determination. Regressors unique to the social capital equation are reflected through Z_t.

[5] We also estimated this relationship as a multinomial logit and linear probability model with not improvement in fit or change in results.

For these identifying (or, excluded) regressors, we construct average census tract characteristics for each observation. The theory is that the characteristics of those in close proximity influence one's level of social capital. Key average characteristics are the share of employment near the person's census tract employed in "social" occupations (as defined in Appendix A) and in social industries which include individuals employed as independent artists, performing arts, and spectator sports (NAICS 856); by bowling centers (858); by religious organizations (916); by civic, social, advocacy organizations and grantmaking services (917); by labor unions (918); and by business, professional, and political organizations (919) (see Rupasingha, Goetz, and Freshwater 2006).

There is some concern that employment levels in civic organizations or social occupations (and other average demographics) *in* a person's census tract may be endogenous to that person's level of social and community involvement. In other words, there are unobservable factors both affecting a person's level of activity and employment in the location where that person has chosen to reside (e.g., someone with high levels of social capital may choose to locate in that census tract because of the high degree of activity by social and community/civic organizations). This potential for endogeneity is why employment levels and other average demographic characteristics in *surrounding* census tracts will be used as the instrument (i.e., excluded regressors), rather than the value of those variables *in* the person's own census tract. These surrounding characteristics will be weighted by the distance (from centroids) of the census tract from that person's census tract. Census tracts in the person's own and surrounding states will be used to construct the average. This method of construction of an instrument in the face of potential geographic endogeneity has been widely applied in the empirical literature (for example, see Y. Lee and Gordon 2005).

2.2 Switching Regression Model of Occupational Wages with Selection

We estimate occupational choice and wages using the 2000 DC data. Wage determination is modeled in two distinct sectors -- social and non-social occupations -- as a switching regression model with known selection (e.g., Willis and Rosen 1979; L.-F. Lee 1978; Idson and Feaster 1990).[6] Because workers make a conscious decision based on the pros and cons of seeking a type of occupation that is classified as a social occupation, wage determination in the labor market can be represented as a three-equation system:

$$W_{s,i} = \beta_s' X_i + \sum_{j=1}^{2} \varphi_{s,j}' \widehat{SK}_{it}^{j} + \varepsilon_{s,i} \qquad if\ SO_i^* > 0\ ; \tag{3}$$

$$W_{ns,i} = \beta_{ns}' X_i + \sum_{j=1}^{2} \varphi_{ns,j}' \widehat{SK}_{it}^{j} + \varepsilon_{ns,i} \quad if\ \ SO_i^* \leq 0\ ; \text{and} \tag{4}$$

$$SO_i^* = \delta E[W_s - W_{ns}]_i + \sum_{j=1}^{2} \alpha_j' \widehat{SK}_{it}^{j} + \gamma' Y_i + u_i\ . \tag{5}$$

$W_{l,i}$ (*l=s,ns*) is the log of hourly wages, X_i are individual characteristics that are expected to influence both wages; β_l (*l=s,ns*) are the returns to measured worker characteristics (subscript *ns* for non-social occupations and *s* for social occupations); and $\widehat{SK}_{it}^{j}$ reflects person *i*'s predicted category *k* for social capital of type *j* (see equation 2), which will be operationalized as dummy variables indicating medium and high levels of each social capital type. Individuals are observed to be employed (or not) in a social occupation based on their latent attraction to that type of occupation, SO_i^*.

Note that we are allowing wages to be directly influenced by a person's level of sociability (as suggested might be important by Deming (2015) and altruism. This influence of social capital characteristics is in addition to any indirect influence those characteristics have on wages through self-selection.

[6] Dolton, Makepeace, and van der Klaauw (1989) illustrate the importance of controlling for self-selection in the determination of occupation-specific wages.

A person's propensity to choose a social occupation (equation 5) is determined by the difference in wages the person expects in choosing a non-social or social occupation, as well as a person's individual social capital, and some characteristics, Y_i, that are expected to affect the decision to choose a social occupation but not determine wages. Variables included in Y_i are the number of children in the household and an indicator for if the person has lived in their current place for five years or less. The person's social capital category *k* (*k*=low, medium, high) of type *j* (*j*=sociability, altruism), $\widehat{SK}_{it}^{j}$, is predicted using parameters obtained from estimating equation 2:

$$\Pr\left[\widehat{SK}_{it}^{j} = k\right] = \frac{1}{\left[1-\exp\left(-\hat{\mu}_{k-1}+\hat{\alpha}_0+\hat{\mathrm{A}}'X_i+\hat{\mathrm{B}}'Y_i+\hat{\mathrm{C}}'Z_t\right)\right]} - \frac{1}{\left[1-\exp\left(-\hat{\mu}_{k}+\hat{\alpha}_0+\hat{\mathrm{A}}'X_i+\hat{\mathrm{B}}'Y_i+\hat{\mathrm{C}}'Z_t\right)\right]},$$

where μ_0 is defined as -∞ and μ_m as +∞.

We would expect that as $E[W_s - W_{ns}]_i$ increases, ceteris paribus, a person would be more likely to choose a social occupation ($\delta > 0$). Since the expected wages in both occupations for each worker are unknown, a reduced form version of equation (5) is estimated by substituting wage determining characteristics, X_i, for the expected wages. $\varepsilon_{l,i}$ (*l*=*ns,s*) and u_i are random error terms that are assumed to be distributed as a trivariate normal. Estimation is performed in multiple stages. We are not modeling selection into the labor market. Consequently, all results will be conditional on a person already having decided to enter the labor market. The joint determination of labor force participation and level of social capital is an interesting question in its own right and will be saved for future research.

2.2.1 Selection into Social and Non-social Occupations. Since a person's social occupation propensity, SO_i^*, is unobserved, equation (5) cannot be directly estimated. Instead, under the assumption of normality the decision of choosing a social occupation can be estimated

via maximum likelihood probit, where a worker is observed in a social occupation if the latent variable $SO_i^* > 0$, and in a non-social occupation if $SO_i^* \leq 0$:

$$\Pr(SocialOcc = 1 | X_i, SK_i, Y_i) = \Phi(\Omega' \mathrm{K}_i),\ \Omega = [\delta, \alpha, \gamma],\ \mathrm{K} = [X_i, SK_i, Y_i]\ . \quad (6)$$

Using the estimated parameter coefficients, inverse mill's ratios are constructed for each observation: $\hat{\lambda}_{s,i} = \frac{\phi(\hat{\Omega}' \mathrm{K}_i)}{\Phi(\hat{\Omega}' \mathrm{K}_i)}$ and $\hat{\lambda}_{ns,i} = -\frac{\phi(\hat{\Omega}' \mathrm{K}_i)}{1-\Phi(\hat{\Omega}' \mathrm{K}_i)}$, where $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal density and cumulative distribution functions, respectively (e.g., see L.-F. Lee 1978)

The inverse mill's ratios are then included as additional regressors in the wage equations such that:

$$E[W_{s,i}] = \beta_s' X_i + \sum_{j=1}^{2} \varphi_{s,j}' \widehat{SK}_{it}^{j} + \theta_s \hat{\lambda}_{s,i} + \varepsilon_{s,i} \qquad for\ SocialOcc = 1 \quad (3')$$

$$E[W_{ns,i}] = \beta_{ns}' X_i + \sum_{j=1}^{2} \varphi_{ns,j}' \widehat{SK}_{it}^{j} + \theta_{ns} \hat{\lambda}_{ns,i} + \varepsilon_{ns,i}\ for\ SocialOcc = 0 \quad (4')$$

Estimation of this specification of the wage equations produces unbiased estimates of the βs, since self-selection into occupational type (social vs. non-social) has been removed from the error term.

It is highly likely that the selection process differs across workers of varying characteristics, and probably most importantly across men and women (see Card 1996). For this reason, and since wage determination has also been found to differ for men and women, we estimate the full model separately by gender.

2.2.2 Decomposition of the Social Occupation Wage Differential. The average observed wage differential between those in social and non-social occupations can be expressed as:

$$\begin{aligned} \overline{W}_s - \overline{W}_{ns} &= \widehat{\Lambda}_s' \overline{\mathrm{M}}_s + \hat{\theta}_s \bar{\hat{\lambda}}_s - \left[\widehat{\Lambda}_{ns}' \overline{\mathrm{M}}_{ns} + \hat{\theta}_{ns} \bar{\hat{\lambda}}_{ns}\right] \\ &= (\overline{\mathrm{M}}_s - \overline{\mathrm{M}}_{ns})' \Lambda^* + \overline{\mathrm{M}}_s' \left(\widehat{\Lambda}_s - \Lambda^*\right) + \overline{\mathrm{M}}_{ns}' \left(\Lambda^* - \widehat{\Lambda}_{ns}\right) + \left[\hat{\theta}_s \bar{\lambda}_s - \hat{\theta}_{ns} \bar{\lambda}_{ns}\right], \end{aligned} \quad (7)$$

where $\overline{\mathrm{M}}_l = [\bar{X}_l, \overline{\widehat{SK}}_{lk}^{c}]$ and $\widehat{\Lambda}_l' = [\hat{\beta}_l', \hat{\varphi}_{l,kc}']$, and *l*=[s,ns]. The first term on the right hand side of the equation is referred to as the endowment effect and indicates how the differences in characteristics of workers in social and non-social occupations contribute to the observed wage differential; the second term is part of what is referred to as the coefficient effect and tells us how wage determination in the non-social occupation world differs from wage determination in some world where endowments in social and non-social occupations were valued equally; the third term tells us how wage determination in social occupations differs from that world with equal valuation of endowments; and the fourth term indicates how differences in selection into the two occupational types influence the differential wages we observe.

Of course, the choice of Λ^* in equation (7) is arbitrary. Several alternatives have been suggested (for example, see Cotton 1988; Reimers 1983). We use a variant on Neumark (1988) and Oaxaca and Ransom (1994) who advocate using a weight matrix to weight both the characteristics and the parameter coefficients, such that:

$$[\overline{W}_s - \overline{W}_{ns}] = (\overline{\mathrm{M}}_s - \overline{\mathrm{M}}_{ns})'\{\boldsymbol{\Psi}\widehat{\Lambda}_s + (\boldsymbol{I} - \boldsymbol{\Psi})\widehat{\Lambda}_{ns}\}$$
$$+\{(\boldsymbol{I} - \boldsymbol{\Psi})'\overline{\mathrm{M}}_s + \boldsymbol{\Psi}'\overline{\mathrm{M}}_{ns}\}(\widehat{\Lambda}_s - \widehat{\Lambda}_{ns}) + [\hat{\theta}_s\bar{\lambda}_s - \hat{\theta}_{ns}\bar{\lambda}_{ns}]. \qquad (7')$$

As suggested by Jann (2008)), we use for $\boldsymbol{\Psi}$ the coefficients from a pooled estimation over both groups, plus an indicator for which group the observation is in. The first term on the right hand side, then, is simply the endowment effect (how much differences in characteristics between those in social and non-social occupations contribute to the wage differential) and the second term is the coefficient effect (how much differences in the valuation of those characteristics in the two occupational sectors contribute to the wage differential).

2.2.3 The Role of Expected Wage Differentials in Occupational Choice. The last step of the analysis is to assess the role of expected wage differentials between social and non-social

occupational outcomes, a person's individual social capital, and other characteristics are the worker's choice of a social occupation. A structural version of the occupational choice model is estimated:

$$SO_i^* = \delta\left[\widehat{W}_s - \widehat{W}_{ns}\right]_i + \sum_{j=1}^{2} \alpha_j' \widehat{SK}_{it}^{j} + \gamma' Y_i + u_i \ . \qquad (5')$$

The coefficient, δ, will tell us how important the expected wage differential is in the person's occupational type choice and the coefficient α will tell us how important a person's individual social capital is in that decision. Because of the nature of the occupational choice -- it's hard to put a dollar value on the utility associated with helping others -- it may very well be the case that a person's expected wage does not play a large role in that choice to pursue a social occupation. For example, Wiswall and Zafar (2016) provide evidence that preferences for a variety of job attributes significantly affect one's occupational choice. It also may be the case that individuals self select such that those whose endowments are highly valued in the non-social occupation world are more likely to choose social occupations, which would produce an estimate of δ that is negative.

2.2.4 Potential Endogeneity of Social Capital. A requirement for the model as detailed above is that social capital is exogenenous to the worker's occupational choice. However, as pointed out by Glaeser, Laibson, and Sacerdote (2002), social capital can be thought of as an investment, undertaken (or enhanced) when positive returns are expected. Consequently, a person in a social occupation may undertake to enhance his/her social capital with the expectation that such activity will be rewarded by the occupation he/she has chosen; Durlauf (2002) warns of various identification issues in the social capital empirical literature. The fact that social capital is not provided in the data set with which we estimate the wage model actually works in our favor here. It is highly unlikely that any of the observations in the SCCBS are the

same respondents completing the one percent sample (long-form) of the decennial census, and because the two samples are from the same population, it is akin to applying split sample IV. Note that the distance-weighted census tract characteristics serve as the instruments for social capital in the occupational choice model and the wage equations.

2.3 Sample Means

Table 1 contains the means of the re-weighted DC sample used to perform the analyses in this paper. The first, perhaps surprising, observation from the means in this table is that the average wage of workers in social occupations is greater than the average wage of workers in non-social occupations. This is not unusual and is consistent with the findings in the literature of a penalty once individual characteristics are controlled for (which is what we find, as well). About 22 percent of women and five percent of men are in social occupations. Those in social occupations are older, more educated, more likely to be married, be a citizen, work for a nonprofit, work for the government and are less likely to be disabled and self-employed. In addition, those in social occupations exhibit higher levels of sociability, but only the highest level of altruism. These differences are fairly consistent across men and women.

[Table 1 about here]

3 Results

The estimation of wages in social and non-social occupations accounting for self-selection is accomplished in a number of stages, as detailed in the previous section. First, the reduced form probability of equation (6) is estimated and inverse mills ratios are constructed; these are relevant to observing someone in a social occupation or not. Second, separate wage equations (depicted in equations 3' and 4') are estimated via OLS and the resulting wage

differentials are decomposed into contributions of endowments of individuals in those occupational groups, the way in which those two markets value those endowments, and the difference in the way individuals select into social and non-social occupations. Third, a structural model containing each person's difference between predicted social occupation and non-social occupation wages, as well as the person's social capital measures is estimated (equation 5). The analysis is performed separately for men and women. All analyses include labor force participants only in order to abstract from decisions to be in the labor force -- hence, all results are to be interpreted as conditional on labor force participation.

3.1 Social Capital Estimation Results

Before any steps of the switching regression model are estimated, estimates of altruism and sociability social capital measures need to be determined for each individual. The methodology for this was described in Section 2.1 above, and the results are found in Table 2.

[Table 2 about here]

Our estimation results link both individual characteristics and the distance weighted census-tract "neighborhood" characteristics with two social capital outcomes. Using the same data set, Brueckner and Largey (2008) estimate two types of social capital -- friendship oriented variables (closest to our sociability) and group-involvement variables (closest to our altruism). Our results are fairly consistent with their findings. For example, they also find that tract density (which will be captured in our analysis by our neighborhood characteristic of the share of population living in MSA - likely highly correlated with population density, which is statistically insignificant) is negatively related to social capital (also see Putnam 1995a); social capital is decreasing at an increasing rate in age and is also lower for married individuals (sociability only), racial (sociability only) and ethnic minorities, and noncitizens; social capital is increasing

in income and education (also see DiPasquale and Glaeser 1999; Helliwell and Putnam 1999; Glaeser, Laibson, and Sacerdote 2002); retired individuals (included in our indicator for not in the labor force) have marginally higher sociability levels; unemployed individuals exhibit lower levels of social capital; someone who has been living in their current location for at least 5 years has higher levels of social capital (also see Schiff 1992; Glaeser et al. 2000; Putnam 1995a).

In addition, Chi squared tests indicate that for each social capital estimation, the group of census tract weighted characteristics of surrounding census tracts (Z_t in equation 2) is statistically significantly different from zero, although individually they have varying degrees of importance in explaining social capital outcomes. Notably, the share of workers in social industries (see Rupasingha, Goetz, and Freshwater 2006) is statistically significantly different from zero in both social capital equations. Recall that the purpose of including these regressors is to be able to identify the impact of the predicted social capital index in the occupational choice equation, independent of the other regressors -- these weighted characteristics are the excluded regressors in the second stage.

Since the parameter estimates from equation (2) will be used to predict levels of social capital out of sample for individuals in the DC, it's important that we have some confidence in the predictive power of the regressors. Table 3 reports, for each actual category, the percent of the sample predicted to be in that category. There is clear dominance in correctly predicting the low and high categories of both social capital measures. However, individuals with actual medium values of each social capital are fairly equally likely to be predicted in the low and high categories as they are to be predicted in the medium category.

[Table 3 about here]

Table 3 also reports the correlation between the actual and predicted social capital categories within the SCCBS sample.[7] The correlations are 0.38 for altruism and 0.30 for sociability, which are quite reasonable compared to the reported fist-stage R-squared statistics obtained by others' applications in an OLS framework of this two-sample prediction strategy. Dee and Evans (2003) make use of predictions from a first stage estimation with R-squared statistics less than 0.02; Carroll, Dynan, and Krane (2003) report first-stage adjusted R-squared test statistics between 0.28 and 0.48; Currie and Yelowitz (2000) report a first-stage R-squared statistic less than 0.1; Fang, Keane, and Silverman (2008) and Keane and Stavrunova (2014) contain similar analyses using the same data and fist-stage R-squared statistics ranging from 0.02 to 0.25. The focus of these authors on their first-stage estimations is the degree to which the excluded regressors (those not included in the second-stage regression) were statistically significantly different from zero (also see Carroll, Dynan, and Krane 2003). Chi squared tests (reported at the bottom of Table 3) indicate that our excluded regressors, distance weighted characteristics of surrounding census tracts, are statistically significantly different from zero.

The social capital indexes estimated and tested here are a combination of answers to various specific questions related to social or civic activities. We also investigated whether the specific, individual activities are more highly correlated with individual and weighted surrounding census tract characteristics. None of the individual responses that we tested provided a better fit of the social capital equation than the indexes reported here.[8]

[7] Since the first stage estimation is an ordered logit, we don't get the usual fit diagnostic of an r-squared test statistic we would get if the first stage were estimated via ordinary least squares.
[8] We tested responses to questions about the amount of time spent watching TV; the number of times in a year the individual gave blood, invited friends to their home, and the amount of time spent volunteering; whether the individual contributed money to religious or other organizations; and how often the individual talked to his/her neighbors.

3.2 Switching Regression Step 1 -- Reduced-form Probit of Social Occupation

Appendix C (Table C1) contains the results from this first-stage, reduced-form probit estimation for men and women separately. The most important observation from these results is that at least one of the identifying regressors are statistically significant determinates of being observed in a social occupation (for both men and women). These are reduced for estimates, so they aren't worth spending too much time on, but for the most part, characteristics of men and women act the same in the determination of being in a social occupation. Some differences include being a high school graduate, some of the regional indicators, Hispanic, marital status, being a citizen, and having lived in the area for five years or less. For example, married women (men) are more (less) likely to be in a social occupation and women (men) who have lived in an area less than five years are less (more) likely to be in a social occupation. In addition, women (men) with higher values of altruism are less (more) likely to be in a social occupation.

3.3 Switching Regression Step 2 -- Wages in Social and non-Social Occupations

Before turning to the switching regression wage estimates, Table 4 presents that standard results presented in the literature (for example, see England, Budig, and Folbre 2002; Hirsch and Manzella 2015) used to identify a wage penalty for workers in social occupations. Note that in the absence of any covariates, workers in social occupations appear to earn a premium -- 15 percent for females and six percent for males. Once other worker characteristics are controlled for, however, that premium turns into a penalty -- two percent for females and 16 percent for males. So even before accounting for selection into social occupations, it is clear that workers with wage enhancing characteristics are more likely to be found in social than in non-social occupations (also see Leete 2001, who illustrates the importance of worker characteristics in

identifying both negative and positive of nonprofit/for-profit wage differentials across different industries).

[Table 4 about here]

Appendix C (Table C2) contains the parameter estimates from estimating each selectivity-corrected wage equations for social and non-social occupations, for both men and women separately. The usual patterns of wage determination are apparent in these regression results. For example, wages are increasing in education, especially for those in social occupations, and they rise with age, at a decreasing rate. One difference of note for both men and women is that Black, non-Hispanics, all else equal, face a wage penalty in non-social occupations, but a wage premium in social occupations. In addition, the self-employed are better off in non-social occupations (as far as wages are concerned, and all else equal).

It's also of interest to note that workers positively select into both social and non-social occupations, but the selection is particularly strong into social occupations -- see the coefficient estimates on the inverse mills ratio regressor ($\theta_s$and θ_{ns} in equations 3' and 4'). This means there is significant and meaningful correlation between unobserved determinants of wages and the likelihood someone chooses, particularly, a social occupation. The correlation is 28 times higher for women in social occupations than for women in non-social occupations; it is about four times higher among men. In addition, we see from these estimates that both high and medium levels of sociability result in a higher wage, which is consistent with the findings of Deming (2015). We also note that sociability reaps a higher reward in social vs. non-social occupations. Higher levels of altruism are also (mostly) valued by employers; the coefficient for medium altruism for women in social occupations, however, is negative.

Table 5 contains the decomposition of the wage differentials between social and non-social occupational groups. The purpose of the decomposition is to break apart the observed wage differential (seen in the means of Table 1 and in the parameter coefficient in the "No Covariates" specification in Table 4) to determine what portion of that differential is explained by (1) differences in observed characteristics between workers in social and non-social occupations, (2) differences in returns to those characteristics (estimated parameters), and (3) difference in individual selection by those in the two types of occupations.

[Table 5 about here]

First of all, we see again that the raw means tell us that women in social occupations earn, on average, about 15 percent higher wages than women in non-social occupation; men earn about 6 percent more. We then see that the differences in endowments are working to elevate wages of those in social occupations, relative to those in non-social occupations (this is seen by the positive contribution of endowments). In other words, characteristics of workers in social occupations are more wage enhancing than the characteristics of workers in non-social occupations. The primary contributor to the difference in endowments is education (for both men and women) -- social occupations attract highly educated workers. Workers in social occupations are also more highly endowed with altruism and sociability social capital characteristics, which, since they are positively rewarded in the labor market, put upward pressure on the observed wage differential.

In spite of the overall downward pressure that differences in coefficients put on the wage differential, education is more highly rewarded in the social occupation world, putting more upward pressure on the observed wage differential, as are the returns to social capital characteristics. In addition, differences in the types of employers and industries in which workers

are employed give workers in social occupations an advantage. Unfortunately, the dominant contributor to observed wage differences is the constant term. This tells us that wages in both occupation groups are determined by a considerable number of other forces than have not been controlled for in these estimations, and that put considerable downward pressure on the wage differential.

The difference in selection plays a role that is equally large to the contribution of coefficients and even larger than the contribution of differences in endowments. Workers select in such a way that increases observed wages in social occupations more than in non-social occupations. Another way to think about this is that there are more unobserved characteristics of people who choose social occupations that make them more productive (have higher wages), in addition to their observed human capital and levels of social capital.

3.4 Switching Regression Step 3 -- Structural Occupational Choice Equation

The model conjectures that individuals choose to be employed in a social occupation or not in part based on where they think they can earn a higher wage. We now estimate equation (5') to see whether expected wages in the two sectors, at an individual level, impact observing an individual in that sector. This equation also includes as regressors the individual measures of social capital, as well as the regressors excluded from the wage equation. Table 6 contains the results from this structural estimation.

[Table 6 about here]

If wages were unimportant to workers' choices of social vs. non-social occupation, then we would conclude that preferences are more important in dictating occupation choice (irrespective of wages), as was found by Weisbrod (1983) in his assessment of lawyers' choices to practice in the nonprofit sector. Also see Dolton, Makepeace, and van der Klaauw (1989) who

find little influence of predicted earnings in an occupation on the choice of that occupation once they control for the graduate's course of study.

The results in Table 6, however, tell a different story. The positive coefficient on the expected wage differentials between social and non-social occupations suggests that expected wages do play a role in a person's choice of working in a social occupation. The effect is much stronger for women. An expected one percent increase in the wage differential from choosing a social vs. non-social occupation increases the probability of choosing a social occupation by 0.20 of a percentage point for women, and by 0.03 of a percentage point for men. Among the employed, 22 percent of women and five percent of men are in social occupations. So, while the influence of expected wage differentials is statistically significant, the effect could be considered small from a practical perspective.

While having a smaller marginal impact than the wage differential on the choice to be in a social occupation, higher levels of altruism and sociability also positively affect that decision. The other regressors have varying degrees of similar effects for men and women. For example, women with more children are more likely to be in a social occupation, whereas the number of children do not affect the decision of men. However, for both men and women, being in an area for less than five years increases the chances of being in a social occupation. This is of interest since the coefficients on this regressor were negative for women and insignificant for men in the first-stage, reduced-form probit, suggesting a relationship between migration decisions and taking advantage of higher returns to endowments for workers in social occupations.

4 Conclusions and Implications

There are several things we learn from the analysis in this paper. The first is that we confirm what others have found in the literature -- once we control for characteristics of workers, those in social occupations earn less that those in non-social occupations. Workers in social occupations possess higher levels of characteristics that are valued in the labor market -- such as, age, education, and sociability (see Deming 2015; Deming and Kahn 2016). In addition, individual selection into social and non-social occupations makes a significant contribution to the observed wage differential; unobserved characteristics of people who choose social occupations make them more productive (have higher wages) in addition to their observed levels of human capital and social capital. This is consistent with others who have concluded that personality and preferences matter a lot in workers sorting into social and non-social occupations (for example, see Wiswall and Zafar 2016; Speer 2017). There is still very much left unexplained about the determination of wages in both occupational groups, however. The difference in the constant terms across wage equations (the really unexplained) wipes out the other positive coefficient effects. We also find that choosing to be in a social occupation has an economic component to it, as well. Someone is significantly more likely to be observed in a social occupation if his/her wage is expected to be higher there than it would be in a non-social occupation.

The main take-away from the results in this paper is two-fold. First, social and caring occupations are not all about women -- both women *and men* choose occupations (including social occupations) best suited to their bundle of characteristics. Second, choosing a social occupation is not all about preferences -- expected wages play a greater role than even a person's level of social capital in choosing between social and non-social occupations.

References

Angrist, Joshua D., and Alan B. Krueger. 1992. "The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples." *Journal of the American Statistical Association* 87 (418): 328–36. doi:10.2307/2290263.

———. 1995. "Split-Sample Instrumental Variables Estimates of the Return to Schooling." *Journal of Business & Economic Statistics* 13 (2): 225–35. doi:10.2307/1392377.

Bayer, Patrick, Stephen Ross, and Giorgio Topa. 2005. "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes." Working Paper 11019. National Bureau of Economic Research. http://www.nber.org/papers/w11019.

Beaman, Lori A. 2012. "Social Networks and the Dynamics of Labour Market Outcomes: Evidence from Refugees Resettled in the U.S." *The Review of Economic Studies* 79 (1): 128–61. doi:10.1093/restud/rdr017.

Bentolila, Samuel, Claudio Michelacci, and Javier Suarez. 2010. "Social Contacts and Occupational Choice." *Economica* 77 (305): 20–45. doi:10.1111/j.1468-0335.2008.00717.x.

Brueckner, Jan K., and Ann G. Largey. 2008. "Social Interaction and Urban Sprawl." *Journal of Urban Economics* 64 (1): 18–34. doi:10.1016/j.jue.2007.08.002.

Card, David. 1996. "The Effect of Unions on the Structure of Wages: A Longitudinal Analysis." *Econometrica* 64 (4): 957–79. doi:10.2307/2171852.

Carroll, Christopher D., Karen E. Dynan, and Spencer D. Krane. 2003. "Unemployment Risk and Precautionary Wealth: Evidence from Households' Balance Sheets." *The Review of Economics and Statistics* 85 (3): 586–604. http://www.jstor.org/stable/3211699.

Cingano, Federico, and Alfonso Rosolia. 2012. "People I Know: Job Search and Social Networks." *Journal of Labor Economics* 30 (2): 291–332. doi:10.1086/663357.

Coleman, James S. 1988. "Social Capital in the Creation of Human Capital." *American Journal of Sociology* 94 (January): S95–120. http://www.jstor.org/stable/2780243.

Cortes, Guido Matias, Nir Jaimovich, and Henry E. Siu. 2016. "The End of Men and Rise of Women in the High-Skilled Labor Market." Mimeo. University of Manchester. https://www.frbatlanta.org/-/media/Documents/news/conferences/2016/0922-unemployment-wages-productivity/papers/jaimovich-end-of-men-rise-of-women-high-skilled-labor-market.pdf.

Cotton, Jeremiah. 1988. "On the Decomposition of Wage Differentials." *The Review of Economics and Statistics* 70 (2): 236–43. doi:10.2307/1928307.

Currie, Janet, and Aaron Yelowitz. 2000. "Are Public Housing Projects Good for Kids?" *Journal of Public Economics* 75 (1): 99–124. doi:10.1016/S0047-2727(99)00065-1.

Dee, Thomas S., and William N. Evans. 2003. "Teen Drinking and Educational Attainment: Evidence from Two-Sample Instrumental Variables Estimates." *Journal of Labor Economics* 21 (1): 178–209. doi:10.1086/344127.

Deller, Steven C., Tsung-Hsiu Tsai, David W. Marcouiller, and Donald B. K. English. 2001. "The Role of Amenities and Quality of Life in Rural Economic Growth." *American Journal of Agricultural Economics* 83 (2): 352–65. http://www.jstor.org/stable/1244678.

Deming, David J. 2015. "The Growing Importance of Social Skills in the Labor Market." Working Paper 21473. National Bureau of Economic Research. http://www.nber.org/papers/w21473.

Deming, David J., and Lisa Kahn. 2016. “Firm Heterogeneity in Skill Demands.” Mimeo. Yale University. http://faculty.som.yale.edu/lisakahn/documents/DemingKahn6_3_16.pdf.
DiNardo, John, Nicole M. Fortin, and Thomas Lemieux. 1996. “Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach.” *Econometrica* 64 (5): 1001–44. doi:10.2307/2171954.
DiPasquale, Denise, and Edward L. Glaeser. 1999. “Incentives and Social Capital: Are Homeowners Better Citizens?” *Journal of Urban Economics* 45 (2): 354–84. doi:10.1006/juec.1998.2098.
Dolton, P. J., G. H. Makepeace, and W. van der Klaauw. 1989. “Occupational Choice and Earnings Determination: The Role of Sample Selection and Non-Pecuniary Factors.” *Oxford Economic Papers*, New Series, 41 (3): 573–94. http://www.jstor.org/stable/2663246.
Durlauf, Steven N. 2002. “On the Empirics of Social Capital.” *The Economic Journal* 112 (483): F459–79. http://www.jstor.org/stable/798457.
England, Paula, Michelle Budig, and Nancy Folbre. 2002. “Wages of Virtue: The Relative Pay of Care Work.” *Social Problems* 49 (4): 455–73. doi:10.1525/sp.2002.49.4.455.
Fang, Hanming, Michael P. Keane, and Dan Silverman. 2008. “Sources of Advantageous Selection: Evidence from the Medigap Insurance Market.” *Journal of Political Economy* 116 (2): 303–3502–303.
Glaeser, Edward L., David I. Laibson, José A. Scheinkman, and Christine L. Soutter. 2000. “Measuring Trust.” *The Quarterly Journal of Economics* 115 (3): 811–46. doi:10.1162/003355300554926.
Glaeser, Edward L., David Laibson, and Bruce Sacerdote. 2002. “An Economic Approach to Social Capital.” *The Economic Journal* 112 (483): F437–58. http://www.jstor.org/stable/798456.
Gould, Eric D. 2002. “Rising Wage Inequality, Comparative Advantage, and the Growing Importance of General Skills in the United States.” *Journal of Labor Economics* 20 (1): 105–47. doi:10.1086/323934.
Hanifan, L. J. 1916. “The Rural School Community Center.” *Annals of the American Academy of Political and Social Science* 67 (September): 130–38. http://www.jstor.org/stable/1013498.
Heckman, James J., and Bo E. Honoré. 1990. “The Empirical Content of the Roy Model.” *Econometrica* 58 (5): 1121–49. doi:10.2307/2938303.
Heckman, James J., and Guilherme L. Sedlacek. 1990. “Self-Selection and the Distribution of Hourly Wages.” *Journal of Labor Economics* 8 (1): S329–63. http://www.jstor.org/stable/2535216.
Heckman, James J., Jora Stixrud, and Sergio Urzua. 2006. “The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior.” *Journal of Labor Economics* 24 (3): 411–82. doi:10.1086/504455.
Hellerstein, Judith K., Melissa McInerney, and David Neumark. 2008. “Neighbors And Co-Workers: The Importance Of Residential Labor Market Networks.” Working Paper 14201. National Bureau of Economic Research. http://www.nber.org/papers/w14201.
Helliwell, John F., and Robert D. Putnam. 1999. “Education and Social Capital.” Working Paper 7121. National Bureau of Economic Research. http://www.nber.org/papers/w7121.

Hensvik, Lena, and Oskar Nordström Skans. 2016. "Social Networks, Employee Selection, and Labor Market Outcomes." *Journal of Labor Economics* 34 (4): 825–67. doi:10.1086/686253.

Hirsch, Barry T., and Julia Manzella. 2015. "Who Cares ? And Does It Matter? Measuring Wage Penalties for Caring Work." In *Gender Convergence in the Labor Market*, 41:213–75. Research in Labor Economics 41. Emerald Group Publishing Limited. http://www.emeraldinsight.com/doi/abs/10.1108/S0147-912120140000041014.

Holzer, Harry J. 1987. "Informal Job Search and Black Youth Unemployment." *The American Economic Review* 77 (3): 446–52. http://www.jstor.org/stable/1804107.

Idson, Todd L., and Daniel J. Feaster. 1990. "A Selectivity Model of Employer-Size Wage Differentials." *Journal of Labor Economics* 8 (1): 99–122. http://www.jstor.org/stable/2535300.

Inoue, Atsushi, and Gary Solon. 2010. "Two-Sample Instrumental Variables Estimators." *Review of Economics and Statistics* 92 (3): 557–61. doi:10.1162/REST_a_00011.

Jann, Ben. 2008. "The Blinder-Oaxaca Decomposition for Linear Regression Models." *The Stata Journal* 8 (4): 453–79. http://www.stata-journal.com/sjpdf.html?articlenum=st0151.

Keane, Michael P, and Olena Stavrunova. 2014. "Adverse Selection, Moral Hazard and the Demand for Medigap Insurance." *Economics Papers from Economics Group,Nuffield College, University of Oxford* No. 2014-W02. http://scor.com/en/sgrc/economy-finance/low-rates-environment/item/3174/3174.html?lout=sgrc.

Kilbourne, Barbara Stanek, Paula England, George Farkas, Kurt Beron, and Dorothea Weir. 1994. "Returns to Skill, Compensating Differentials, and Gender Bias: Effects of Occupational Characteristics on the Wages of White Women and Men." *American Journal of Sociology* 100 (3): 689–719. http://www.jstor.org/stable/2782402.

Kuhn, Peter, and Catherine Weinberger. 2005. "Leadership Skills and Wages." *Journal of Labor Economics* 23 (3): 395–436. doi:10.1086/430282.

Lee, Lung-Fei. 1978. "Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables." *International Economic Review* 19 (2): 415–33. doi:10.2307/2526310.

Lee, Young, and Roger H. Gordon. 2005. "Tax Structure and Economic Growth." *Journal of Public Economics* 89 (5–6): 1027–43. doi:10.1016/j.jpubeco.2004.07.002.

Leete, Laura. 2001. "Whither the Nonprofit Wage Differential? Estimates from the 1990 Census." *Journal of Labor Economics* 19 (1): 136–70. doi:10.1086/209982.

Lin, Nan. 1999. "Social Networks and Status Attainment." *Annual Review of Sociology* 25 (January): 467–87. http://www.jstor.org/stable/223513.

Mani, Anandi, and Charles H. Mullin. 2004. "Choosing the Right Pond: Social Approval and Occupational Choice." *Journal of Labor Economics* 22 (4): 835–61. doi:10.1086/423156.

Montgomery, James D. 1991. "Social Networks and Labor-Market Outcomes: Toward an Economic Analysis." *The American Economic Review* 81 (5): 1408–18. http://www.jstor.org/stable/2006929.

Mouw, Ted. 2003. "Social Capital and Finding a Job: Do Contacts Matter?" *American Sociological Review* 68 (6): 868–98. doi:10.2307/1519749.

Neumark, David. 1988. "Employers' Discriminatory Behavior and the Estimation of Wage Discrimination." *The Journal of Human Resources* 23 (3): 279. doi:10.2307/145830.

Ngom, Pappe. 2003. “The Role of Individual Social Capital in Occupational Choice and Earnings.” *Doctoral Dissertations*, January, 1–117. http://digitalcommons.uconn.edu/dissertations/AAI3116837.

Oaxaca, Ronald, and Michael Ransom. 1994. “On Discrimination and the Decomposition of Wage Differentials.” *Journal of Econometrics* 61 (1): 5–21. http://econpapers.repec.org/article/eeeeconom/v_3a61_3ay_3a1994_3ai_3a1_3ap_3a5-21.htm.

Pitts, M. Melinda. 2003. “Why Choose Women’s Work If It Pays Less? A Structural Model of Occupational Choice.” In *Worker Well-Being and Public Policy*, 22:415–45. Research in Labor Economics 22. Emerald Group Publishing Limited. http://www.emeraldinsight.com/doi/abs/10.1016/S0147-9121%2803%2922013-6.

Putnam, Robert D. 1994. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton University Press.

———. 1995a. *Bowling Alone: America’s Declining Social Capital*. Simon & Schuster.

———. 1995b. “Tuning In, Tuning Out: The Strange Disappearance of Social Capital in America.” *PS: Political Science & Politics* 28 (4): 664–683. doi:10.2307/420517.

Reimers, Cordelia W. 1983. “Labor Market Discrimination against Hispanic and Black Men.” *The Review of Economics and Statistics* 65 (4): 570–79. http://econpapers.repec.org/article/tprrestat/v_3a65_3ay_3a1983_3ai_3a4_3ap_3a570-79.htm.

Rencher, Alvin C. 1997. *Multivariate Statistical Inference and Applications*. 1 edition. New York: Wiley-Interscience.

Ridder, Geert, and Robert Moffitt. 2007. “Chapter 75 The Econometrics of Data Combination.” In *Handbook of Econometrics*, edited by James J. Heckman and Edward E. Leamer, 6, Part B:5469–5547. Elsevier. http://www.sciencedirect.com/science/article/pii/S1573441207060758.

Roy, A. D. 1951. “Some Thoughts on the Distribution of Earnings.” *Oxford Economic Papers*, New Series, 3 (2): 135–46. http://www.jstor.org/stable/2662082.

Rupasingha, Anil, Stephan J. Goetz, and David Freshwater. 2006. “The Production of Social Capital in US Counties.” *The Journal of Socio-Economics*, Essays on Behavioral Economics, 35 (1): 83–101. doi:10.1016/j.socec.2005.11.001.

Saffer, Henry, and Karine Lamiraud. 2008. “The Effect of Hours of Work on Social Interaction.” Working Paper 13743. National Bureau of Economic Research. http://www.nber.org/papers/w13743.

Schiff, M. 1992. “Social Capital, Labor Mobility, and Welfare: The Impact of Uniting States.” *Rationality and Society* 4 (2): 157–75. doi:10.1177/1043463192004002003.

Schmutte, Ian M. 2015. “Job Referral Networks and the Determination of Earnings in Local Labor Markets.” *Journal of Labor Economics* 33 (1): 1–32. doi:10.1086/677389.

Speer, Jamin D. 2017. “Pre-Market Skills, Occupational Choice, and Career Progression.” *Journal of Human Resources* Forthcoming. https://www.dropbox.com/s/8oybzam40vsqz82/Speer_Occupations_July2015.pdf?dl=0.

Wegener, Bernd. 1991. “Job Mobility and Social Ties: Social Resources, Prior Job, and Status Attainment.” *American Sociological Review* 56 (1): 60–71. doi:10.2307/2095673.

Weisbrod, Burton A. 1983. “Nonprofit and Proprietary Sector Behavior: Wage Differentials among Lawyers.” *Journal of Labor Economics* 1 (3): 246–63. http://www.jstor.org/stable/2534835.

Willis, Robert J., and Sherwin Rosen. 1979. “Education and Self-Selection.” *Journal of Political Economy* 87 (5): S7–36. http://www.jstor.org/stable/1829907.
Wiswall, Matthew, and Basit Zafar. 2016. “Preference for the Workplace, Human Capital, and Gender.” Working Paper 22173. National Bureau of Economic Research. http://www.nber.org/papers/w22173.
Wollcock, M. 2001. “The Place of Social Capital in Understanding Social and Economic Outcomes.” *Canadian Journal of Public Policy* 2: 11–17.

Table 1. Sample means of 2000 decennial census.

	Female		Male	
Variable	Social Occupations	Non-social Occupations	Social Occupations	Non-social Occupations
Log hourly wage	2.5769	2.4247	2.755	2.6997
	(.7084)	(.691)	(.7144)	(.7551)
Less than high school education=0,1	.0671	.133	.0392	.17
	(.2502)	(.3396)	(.1942)	(.3757)
High school education=0,1	.1506	.3069	.0765	.2875
	(.3577)	(.4612)	(.2658)	(.4526)
Some college education=0,1	.2881	.3595	.2023	.2987
	(.4529)	(.4799)	(.4017)	(.4577)
College graduate=0,1	.4942	.2006	.6819	.2438
	(.5)	(.4004)	(.4657)	(.4293)
White, non-Hispanic=0,1	.736	.71	.7477	.7218
	(.4408)	(.4537)	(.4343)	(.4481)
Hispanic=0,1	.0773	.1071	.073	.1281
	(.267)	(.3093)	(.2601)	(.3342)
Black, non-Hispanic=0,1	.1372	.1223	.1143	.0921
	(.344)	(.3276)	(.3181)	(.2891)
Other race, non-hispanic=0,1	.0496	.0605	.0651	.0581
	(.2171)	(.2385)	(.2467)	(.2339)
Live in MSA=0,1	.8012	.8194	.8167	.8202
	(.3991)	(.3847)	(.3869)	(.384)
Age	40.9762	39.3914	42.251	39.8115
	(12.5152)	(13.1238)	(13.715)	(13.1771)
Married=0,1	.6191	.5477	.6394	.6178
	(.4856)	(.4977)	(.4802)	(.4859)
Citizen=0,1	.955	.935	.9375	.9079
	(.2073)	(.2466)	(.2421)	(.2892)
disability=0,1	.136	.1533	.133	.1686
	(.3428)	(.3603)	(.3396)	(.3744)

	Female		Male	
Variable	Social Occupations	Non-social Occupations	Social Occupations	Non-social Occupations
Private for-profit employer=0,1	.3872	.7719	.229	.7986
	(.4871)	(.4196)	(.4202)	(.401)
Private nonprofit employer=0,1	.224	.0732	.2848	.0317
	(.4169)	(.2605)	(.4513)	(.1753)
Government employer=0,1	.3746	.1256	.4734	.1097
	(.484)	(.3314)	(.4993)	(.3125)
Self-employed=0,1	.0141	.0293	.0128	.06
	(.1181)	(.1686)	(.1123)	(.2374)
altruism = low	.1654	.2963	.1915	.4333
	(.3715)	(.4566)	(.3935)	(.4955)
altruism = medium	.2935	.3868	.2857	.322
	(.4554)	(.487)	(.4517)	(.4672)
altruism = high	.5411	.3169	.5228	.2447
	(.4983)	(.4653)	(.4995)	(.4299)
sociability = low	.3779	.5398	.1796	.414
	(.4849)	(.4984)	(.3838)	(.4925)
sociability = medium	.3437	.2467	.2452	.2404
	(.475)	(.4311)	(.4302)	(.4273)
sociability = high	.2783	.2134	.5753	.3456
	(.4482)	(.4097)	(.4943)	(.4756)
Number of children in HH	1.0122	1.0041	.9372	1.0203
	(1.177)	(1.1788)	(1.1686)	(1.2015)
Household total income GE $30,000=0,1	.7926	.7676	.8155	.7963
	(.4055)	(.4224)	(.3879)	(.4027)
Own home=0,1	.7457	.7101	.7269	.7313
	(.4355)	(.4537)	(.4456)	(.4433)
Lived in area 5 yrs or less=0,1	.298	.3251	.3459	.3276
	(.4574)	(.4684)	(.4757)	(.4693)
Observations	2,150,000	7,740,000	560,000	9,880,000

Notes: Standard deviations in parentheses. Sample sizes are rounded to the nearest 10,000 for disclosure purposes.

Table 2. Ordered logit parameter estimates of determinants of individual social capital measures

Individual Characteristics	Altruism	Sociability
Age	-0.1243*	-0.2277***
	(0.0681)	(0.0684)
Age squared	0.0036**	0.0040***
	(0.0015)	(0.0016)
Married=0,1	0.0713	-0.1502***
	(0.0551)	(0.0543)
Number of children in HH	0.2448***	0.0231
	(0.0211)	(0.0208)
Household total income GE $30,000=0,1	0.4599***	0.4592***
	(0.0614)	(0.0621)
High school education=0,1	0.2849**	0.2640**
	(0.1187)	(0.1238)
Some college education=0,1	0.7949***	0.7864***
	(0.1198)	(0.1262)
College graduate=0,1	1.3739***	1.1122***
	(0.1643)	(0.1711)
Hispanic=0,1	-0.2527**	-0.4723***
	(0.1270)	(0.1263)
Black, non-Hispanic=0,1	0.0369	-0.6322***
	(0.1112)	(0.1123)
Other race, non-hispanic=0,1	-0.2729	-0.9840***
	(0.2028)	(0.1808)
Unemployed=0,1	-0.4022***	-0.3757***
	(0.1128)	(0.1133)
Not in the labor force=0,1	0.0451	0.1214*
	(0.0648)	(0.0664)
Citizen=0,1	0.5866***	0.3252***
	(0.1253)	(0.1193)
Lived in area 5 yrs or less=0,1	-0.2217***	-0.3095***
	(0.0529)	(0.0497)

Individual Characteristics	Altruism	Sociability
Own home=0,1	0.3273***	0.2982***
	(0.1040)	(0.1108)
Female=0,1	0.2505***	-0.2370***
	(0.0450)	(0.0459)
Live in MSA=0,1	0.0941	-0.0582
	(0.0930)	(0.0851)
Mid Atlantic region=0,1	-0.0614	0.1273
	(0.1664)	(0.1650)
East North Central region=0,1	0.1471	0.3431*
	(0.1882)	(0.1889)
West North Central region=0,1	0.1756	0.2154
	(0.2583)	(0.2618)
South Atlantic region=0,1	0.2842	0.2091
	(0.2222)	(0.2194)
East South Central region=0,1	0.1466	0.2911
	(0.2716)	(0.2597)
West South Central region=0,1	0.5424**	0.5888**
	(0.2575)	(0.2616)
Mountain region=0,1	0.3185	0.5342**
	(0.2397)	(0.2425)
Pacific region=0,1	0.1398	0.4567**
	(0.2255)	(0.2271)
College grad * white non-Hispanic	-0.0356	0.0425
	(0.1205)	(0.1263)
College grad * own home	0.1324	-0.1262
	(0.0992)	(0.1046)
White non-Hispanic * own home	-0.0727	-0.4037***
	(0.1186)	(0.1176)
Age GE 75 years	1.2507**	0.7380
	(0.6014)	(0.6558)

Individual Characteristics	Altruism	Sociability
Age LT 25 years	0.1861	0.2350
	(0.1354)	(0.1443)
Age cubed (0000)	-0.2930***	-0.2268**
	(0.1114)	(0.1145)
Distance-weighted Surrounding Census Tract Characteristics[a]		
Share of workers in broad SK occupations	7.4662	2.9405
	(5.1171)	(5.3932)
Share of workers in SK industries	47.7027***	26.7807*
	(14.4044)	(15.3310)
Labor force participation rate	8.5483	2.9756
	(8.9398)	(8.4981)
Unemployment rate	24.9550	39.9777
	(26.5612)	(25.4096)
Percent lived in area at least 5 years	-6.2586***	-2.0281
	(1.7620)	(1.6643)
Median age	0.0756	0.0683
	(0.0509)	(0.0527)
Diversity index	0.5054	-0.4698
	(0.8512)	(0.8790)
Female labor force participation rate	-6.8299	-0.7531
	(8.1924)	(7.6405)
Percent college graduates, 25 and older	-4.8330**	0.0926
	(2.4200)	(2.3454)
Percent married households	-4.2712**	-2.3004
	(2.1506)	(2.0793)
Percent of families with children	-2.0880	-0.8611
	(3.0068)	(3.1106)
Percent who own home	1.5886	0.7356
	(1.6166)	(1.5586)
Median household income ($00000)	2.8267*	0.5578
	(1.6663)	(1.7441)

Individual Characteristics	Altruism	Sociability
Population density (1000/sq mi)	0.0250	0.0280
	(0.0455)	(0.0502)
Share of population living in MSA	-1.0102**	-0.9843**
	(0.4152)	(0.4317)
Constant cut1	0.6801	-0.9932
	(3.7104)	(3.7618)
Constant cut2	2.3568	0.5467
	(3.7121)	(3.7611)
Observations	18,716	18,716
Chi2, census-weighted demographic parameters all zero	45.20	39.94

Notes: Data used for analysis are those from the Social Capital Benchmark Survey. Dependent variable are Altruism or Sociability = 0,1,2. Observations are weighted using an inverse probability adjustment to the weights supplied by the SCCBS (see DiNardo, Fortin, and Lemieux 1996). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

[a]These regressors are unique to the social capital prediction equation.

Table 3. Actual versus predicted categories of social capital measures.

	Altruism	Sociability
Actual Level of Low		
Predicted level of Low	52.94%	57.26%
Predicted level of Medium	33.54%	23.44%
Predicted level of High	13.52%	19.30%
Actual Level of Medium		
Predicted level of Low	32.27%	38.82%
Predicted level of Medium	38.08%	28.41%
Predicted level of High	29.65%	32.77%
Actual Level of High		
Predicted level of Low	15.24%	24.35%
Predicted level of Medium	33.60%	26.21%
Predicted level of High	51.16%	49.44%
Correlations between actual and predicted categories across individuals	0.38	0.30

Table 4. OLS wage regressions without selection.

	Females		Males	
	No Covariates	Including Covariates	No Covariates	Including Covariates
Coefficient on social occupation indicator = (0.1)	0.1522***	-0.0223***	0.0553***	-0.1552***
	(0.0006)	(0.0007)	(0.0011)	(0.0012)
Observations	9,880,000		10,440,000	

Notes: Additional covariates include the usual demographic and human capital characteristics, found listed in Table C2 in Appendix C. Sample sizes are rounded to the nearest 10,000 for disclosure purposes.

Table 5. Decomposition of wage differentials between workers in social and non-social occupations, separately for men and women.

		Females	Males
Average log wage social occupations		2.577***	2.755***
		(0.0005)	(0.0011)
Average log wage non-social occupations		2.425***	2.700***
		(0.0003)	(0.0003)
Observed log wage differential	$[\overline{W}_s - \overline{W}_{ns}] =$	0.1522***	0.0553***
		(0.0006)	(0.0011)
Contribution of differences in:			
Endowments	$(\overline{\mathrm{M}}_s - \overline{\mathrm{M}}_{ns})'\{\boldsymbol{\Psi}\widehat{\Lambda}_s + (\boldsymbol{I} - \boldsymbol{\Psi})\widehat{\Lambda}_{ns}\} =$	**0.1373*** **	**0.2482*** **
		(0.0016)	(0.0015)
education characteristics		0.1232***	0.1689***
		(0.0003)	(0.0005)
race and ethnicity characteristics		0.0010***	0.0031***
		(0.0000)	(0.0001)
region dummies, plus MSA		-0.0015***	0.0023***
		(0.0001)	(0.0001)
other demographics (age, marital status, citizen, disability)		0.0252***	0.0292***
		(0.0001)	(0.0003)
job characteristics (type of employer, industries dummies)		-0.0335***	-0.0017
		(0.0014)	(0.0013)
social capital characteristics (altruism and sociability)		0.0229***	0.0464***
		(0.0002)	(0.0003)
Coefficients	$\{(\boldsymbol{I} - \boldsymbol{\Psi})'\overline{\mathrm{M}}_s + \boldsymbol{\Psi}'\overline{\mathrm{M}}_{ns}\}(\widehat{\Lambda}_s - \widehat{\Lambda}_{ns}) =$	**-0.8192*** **	**-1.134*** **
		(0.0111)	(0.0300)
education coefficients		0.1988***	0.3422***
		(0.0035)	(0.0162)
race and ethnicity coefficients		0.0193***	0.0375***
		(0.0004)	(0.0008)
region coefficients, plus MSA		-0.1309***	-0.0896***
		(0.0025)	(0.0045)
other demographics (age, marital status, citizen, disability		-0.2335***	-0.3862***
		(0.0086)	(0.0196)
job coefficients (type of employer, industries dummies)		1.548***	1.405***
		(0.0184)	(0.0430)
social capital coefficients (altruism and sociability)		0.0161***	0.0161***
		(0.0018)	(0.0038)
Constant		-2.237***	-2.459***
		(0.0285)	(0.0738)
Selection	$[\hat{\theta}_s\bar{\lambda}_s - \hat{\theta}_{ns}\bar{\lambda}_{ns}] =$	**0.8341*** **	**0.9410*** **
		(0.0104)	(0.0298)

	Females	Males
Observations	9,880,000	10,440,000

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample sizes are rounded to the nearest 10,000 for disclosure purposes. Social occupations are defined in Appendix A.

Table 6. Structural probit estimation of occupational choice with expected individual wage differential.

Variables	Female	Male
Expected log wage differential $\left[\widehat{W}_s - \widehat{W}_{ns}\right]_i$	**1.208*****	**1.496*****
	(0.0009)	(0.0017)
	[0.1990]	[0.0267]
altruism = medium	0.0053***	0.0793***
	(0.0019)	(0.0030)
	[0.0009]	[0.0015]
altruism = high	0.0270***	0.1556***
	(0.0020)	(0.0031)
	[0.0045]	[0.0031]
sociability = medium	0.0046***	0.0864***
	(0.0017)	(0.0033)
	[0.0007]	[0.0016]
sociability = high	0.0334***	0.0938***
	(0.0018)	(0.0031)
	[0.0056]	[0.0017]
Number of children in HH	0.0187***	0.0003
	(0.0006)	(0.0009)
	[0.0031]	[0.0000]
Lived in area 5 yrs or less=0,1	0.0193***	0.0979***
	(0.0015)	(0.0023)
	[0.0032]	[0.0018]
Constant	1.143***	1.219***
	(0.0021)	(0.0040)
Observations	9,880,000	10,440,000

Notes: Standard errors in parentheses; marginal effects in brackets (calculated as discrete change for dummy variables). *** p<0.01, ** p<0.05, * p<0.1. Sample sizes are rounded to the nearest 10,000 for disclosure purposes. The average predicted wage differential (roughly interpreted as the social occupation wage "penalty") between social and non-social occupations is -2.09 (s.d. 0.99) for women and -2.58 (s.d. 0.65) for men.

Appendix A. Social Occupations

A.1 Classifying Social Occupations

We make use of a third-party classification of what occupations are to be considered "social." These occupations overlap what others have classified (using varying classification techniques) as sociable and/or caring. The classification that we used (obtained from the career counseling website, Career Key, notably excludes high-end/entrepreneurial occupations, such as doctor or lawyer (occupations that have been included in others' classifications). Occupations are listed in Table A1 along with their 2000 Census occupation codes (see https://usa.ipums.org/usa/volii/00occup.shtml).

[Table A1 about here]

*A.2 Validation using O*NET*

The Occupational Information Network (O*NET) database is the central project of the O*NET program developed under the auspices of the US Department of Labor/Employment and Training Administration to serve as the primary source of occupational information in the U.S. The O*NET Content Model provides the conceptual framework for identifying the key measurable features of an occupation called “descriptors”. The model is categorized into 6 broad domains, namely Worker Characteristics, Worker Requirements, Experience Requirements, Occupational Requirements, Workforce Characteristics, and Occupation-Specific Information. The current O*NET 20.2 database has 277 descriptors (http://www.onetcenter.org/content.html) on 953 occupations (http://www.onetcenter.org/taxonomy/2010/updated.html). However, it is the 2000 O*NET 3.0 database that is relevant for our analyses because it coincides with the 2000

Decennial Census.[9] Both databases use the 2000 Standard Occupational Classification (SOC) system to classify occupations.[10]

The O*NET database is continually updated by the O*NET Data Collection Program that administers standardized questionnaires to a random sample of workers in occupations that are part of a statistically random sample of businesses that were selected in a first stage. Based on the responses by the workers to the O*NET questionnaire, several occupational ratings were created and are available in the O*NET database.

Respondents to the O*NET surveys are not asked to provide information for all descriptors as this will be burdensome. Rather, respondents are randomly assigned to one of three questionnaires relating to different data elements. As a result, information is not available for some features or attributes of our social occupations. In particular, the 2000 O*NET database contains hundreds of descriptors grouped into 8 categories: Ability, Skills, Interest, Knowledge, Tasks, Work Activity, Work Context, and Work Value. However, for our social occupations, data was only collected on variables within the Work Interest and Work Value categories.[11]

The Work Interest category describes worker preferences for work environments and outcomes. There are 6 attributes or variables used to describe the work environment, namely Realistic, Investigative, Artistic, Social, Enterprising, and Conventional. The Work Values

[9] The 2000 O*NET 3.0 database can be found at http://www.xwalkcenter.org/index.php/component/content/article/83-onetinfo/102-onet3 with corresponding occupation classifications at https://www.onetcenter.org/taxonomy/2000/list.html.
[10] A crosswalk from the 2000 SOC to the 2000 DC is available at http://www.xwalkcenter.org/index.php/classifications/crosswalks.
[11] Some variables in the O*NET database have multiple measures or scales. For instance, variables within the Ability category are measured on "importance" and "level" scales. Previous authors creating indices using O*NET have combined these different measures or scales by assigning Cobb-Douglas weights to the different measures or scales (e.g., see Blinder 2007 and Firpo et al. 2011). This type of reconstruction is not relevant for our analyses because the variables we use from the Interest and Work Value are measured on a single scale or dimension, namely, the "occupational interest" and "extent" scales respectively.

category describes features or aspects of work that are important to a person's satisfaction. These features are Achievement, Working Conditions, Recognition, Relationships, Support, and Independence.

Table A2 compares the mean attribute values for Worker Occupational Interests and Work Values for all occupations in the O*NET data set with the values specific to the occupations listed in Table A1 as social occupations. Comparing the means in the table, there are a number of ways in which social occupations stand out in expected ways from the sample of all occupations. The values and interests of workers in social occupations that are *lower*, at least at the 95 percent confidence level, than workers in all occupations include: Realistic (working with real world materials such as machines, plants, animals, etc.), Compensation (paid well compared to other jobs), and Independence (working alone). The values and interests of workers in social occupations that *exceed* those of workers in all occupations include: Artistic (involving self-expression without a clear set of rules), Social (working and communicating with and often helping others), Authority (gives directions and instructions to others), and Social Service (do things for other people). The point of this exercise is to demonstrate that the occupations that we use as social occupations are characteristically those in line with altruistic and sociability attributes.

[Table A2 about here]

Table A1. Occupations classified as social occupations.

2000 Occupation Codes	Occupation Title
Community and Social Services Occupations	
200	Counselors
201	Social Workers
202	Miscellaneous Community and Social Service Specialists
203	Not used
204	Clergy
205	Directors, Religious Activities and Education
206	Religious Workers, All Other
Select Education, Training, and Library Occupations	
220	Postsecondary Teachers
230	Preschool and Kindergarten Teachers
231	Elementary and Middle School Teachers
232	Secondary School Teachers
233	Special Education Teachers
234	Other Teachers and Instructors
243	Librarians
244	Library Technicians
254	Teacher Assistants
255	Other Education, Training, and Library Workers
Select Healthcare Practitioners and Technical Occupations	
303	Dietitians and Nutritionists
311	Physician Assistants
313	Registered Nurses
314	Audiologists
315	Occupational Therapists
316	Physical Therapists
320	Radiation Therapists

2000 Occupation Codes	Occupation Title
321	Recreational Therapists
322	Respiratory Therapists
323	Speech-Language Pathologists
324	Therapists, All Other
331	Dental Hygienists
340	Emergency Medical Technicians and Paramedics
Healthcare Support Occupations	
360	Nursing, Psychiatric, and Home Health Aides
361	Occupational Therapist Assistants and Aides
362	Physical Therapist Assistants and Aides
363	Massage Therapists
364	Dental Assistants
365	Medical Assistants and Other Healthcare Support Occupations
Select Protective Service Occupations	
394	Crossing Guards
395	Lifeguards and Other Protective Service Workers
Select Personal Care and Service Occupations	
460	Child Care Workers
461	Personal and Home Care Aides
462	Recreation and Fitness Workers
465	Personal Care and Service Workers, All Other

Notes. Occupation codes obtained found at https://usa.ipums.org/usa/volii/00occup.shtml. Occupation classification of social occupations obtained from Career Key, an online career services organization (https://www.careerkey.org/explore-career-options/social-careers-career-clusters.html). Occupations not identifiable by 2000 Occupational Codes include "Coach or Scout Leader," "Sports Official or Umpire," "Recreation Worker," and "Fitness Trainer."

Table A2. Comparison of average values of occupation attributes of social occupations with other occupations, using O*NET data

Element Name	Description	All Occupations	Social Occupations
Worker Interest Attributes			
Artistic	Artistic occupations frequently involve working with forms, designs and patterns. They often require self-expression and the work can be done without following a clear set of rules.	2.5619 (1.2295)	3.7430*** (0.9770)
Conventional	Conventional occupations frequently involve following set procedures and routines. These occupations can include working with data and details more than with ideas. Usually there is a clear line of authority to follow.	3.8560 (1.1316)	3.4795*** (0.7358)
Enterprising	Enterprising occupations frequently involve starting up and carrying out projects. These occupations can involve leading people and making many decisions. Sometimes they require risk taking and often deal with business.	3.4190 (1.4281)	3.0810*** (0.6676)
Investigative	Investigative occupations frequently involve working with ideas, and require an extensive amount of thinking. These occupations can involve searching for facts and figuring out problems mentally.	3.1915 (1.4481)	4.1973*** (1.2482)
Realistic	Realistic occupations frequently involve work activities that include practical, hands-on problems and solutions. They often deal with plants, animals, and real-world materials like wood, tools, and machinery. Many of the occupations require working and do not involve a lot of paperwork or working closely with others.	5.1574 (1.5821)	3.6720*** (1.1137)
Social	Social occupations frequently involve working with, communicating with, and teaching people. These occupations often involve helping or providing service to others.	3.0478 (1.5694)	6.0514*** (0.9322)
Work Value Attributes			
Achievement: Occupations that satisfy this work value are results oriented and allow employees to use their strongest abilities, giving them a feeling of accomplishment. Corresponding needs are Ability Utilization and Achievement.			
Ability Utilization	Workers on this job make use of their individual abilities.	3.8581 (0.6899)	3.8653 (0.6366)
Achievement	Workers on this job get a feeling of accomplishment.	3.8164 (0.6429)	4.0937*** (0.5691)

Element Name	Description	All Occupations	Social Occupations
Working Conditions: Occupations that satisfy this work value offer job security and good working conditions. Corresponding needs are Activity, Compensation, Independence, Security, Variety and Working Conditions.			
Activity	Workers on this job are busy all the time	3.4882 (0.4168)	3.5303 (0.3665)
Compensation	Workers on this job are paid well in comparison with other workers	3.2750 (0.5378)	2.9796*** (0.4003)
Independence	Workers on this job do their work alone	2.9087 (0.6048)	2.4196*** (0.3714)
Security	Workers on this job have steady employment	3.6364 (0.4584)	3.6879 (0.2918)
Variety	Workers on this job have something different to do every day	3.2467 (0.4309)	3.2976 (0.2761)
Working Conditions	Workers on this job have good working conditions	3.6170 (0.5934)	3.6851 (0.5916)
Recognition: Occupations that satisfy this work value offer advancement, potential for leadership, and are often considered prestigious. Corresponding needs are Advancement, Authority, Recognition and Social Status.			
Advancement	Workers on this job have opportunities for advancement	2.8528 (0.4514)	2.7396** (0.2889)
Authority	Workers on this job give directions and instructions to others	3.1691 (0.8635)	3.6613*** (0.8124)
Recognition	Workers on this job receive recognition for the work they do	3.2487 (0.6234)	3.1621 (0.5368)
Social Status	Workers on this job are looked up to by others in their company and their community	3.4431 (0.6161)	3.5592 (0.5499)
Independence: Occupations that satisfy this work value allow employs to work on their own and make decisions. Corresponding needs are Creativity, Responsibility and Autonomy.			
Autonomy	Workers on this job plan their work with little supervision	3.7022 (0.7487)	3.5574 (0.8077)
Creativity	Workers on this job try out their own ideas	3.3659 (0.8730)	3.4242 (0.8791)
Responsibility	Workers on this job make decisions on their own	3.6555	3.5582

Element Name	Description	All Occupations	Social Occupations
		(0.6881)	(0.8221)
Relationships: Occupations that satisfy this work value allow employees to provide service to others and work with co-workers in a friendly non-competitive environment. Corresponding needs are Co-workers, Moral Values and Social Service.			
Co-workers	Workers on this job have co-workers who are easy to get along with	3.3196 (0.4919)	3.6695*** (0.4820)
Moral Values	Workers on this job are never pressured to do things that go against their sense of right and wrong	3.5149 (0.4553)	3.5391 (0.3679)
Social Service	Workers on this job have work where they do things for other people	3.0426 (0.9924)	4.2154*** (0.5395)
Support: Occupations that satisfy this work value offer supportive management that stands behind employees. Corresponding needs are Company Policies, Supervision: Human Relations and Supervision: Technical.			
Supervision, Human Relations	Workers on this job have supervisors who back up their workers with management	3.1334 (0.5401)	3.2186* (0.3309)
Supervision, Technical	Workers on this job have supervisors who train their workers well	2.3758 (0.5607)	2.3007 (0.4619)
Company Policies and Practices	Workers on this job are treated fairly by the company	3.3916 (0.5198)	3.4757 (0.4026)

Note: *, **, *** => mean attribute for social occupations differs from the mean value for all occupations at the 90, 95, and 99 percent confidence level, based on a standard Z-statistic.

Appendix B. Factor Analysis to Identify Latent Altruism and Sociability Characteristics

Factor analysis is a data reduction technique for expressing observed variables as linear combinations of a few unobserved variables called *factors.* We use factor analysis to construct 2 factors that we hope capture our underlying social capital dimensions of "sociability" and "altruism" using a multitude of questions from the SCCBS.

Our approach in using factor analysis is both exploratory and confirmatory. It is confirmatory in the sense that we determined *a priori* the two dimensions of social capital that are relevant for our analysis as well as the possible questions from the SCCBS that belong to each dimension. It is exploratory in that we conduct factor analysis within these two categories and retain the first factor extracted. It is typical for factor analysis to produce as many factors as the number of original variables in a particular category. As such, the determination of how many factors to retain from the factor analysis for subsequent analyses is based on well-known rules of thumb (e.g., see (Rencher 1997). One is to retain all factors with eigenvalues greater than one. Another is to choose the number of factors required to reach a given percentage of explained total variation (measured as trace of correlation matrix). Other researchers use statistical hypothesis testing when the maximum likelihood method of estimation of the loadings matrix is adopted, while others simply retain the first factor extracted (e.g., see (Deller et al. 2001). We followed the last rule of thumb in this paper, since a driving motivation for performing factor analysis in this case was for the purposes of data reduction.

For both sociability and altruism dimensions of social capital, we perform factor analysis using the principal component method on the polychoric correlation matrix since the variables in our data are binary, ordinal, or continuous. The tables below summarize results from our factor analysis. The question with the highest factor loading for altruism is whether the person worked

on a community project in the past 12 months, followed by whether the person participates in a social or welfare organization. The question with the highest factor loading for sociability is the question about whether the individual participates in a literary, art, or musical group, followed by a question about whether the person participates in a hobby, investment, or garden club.

Table B1. Factor analysis results for altruism social capital factor.

Variable from SCCBS	Description	Factor Loadings	Correlation between retained factor and original variables	Uniqueness
GRPPTA	Participate in parent association or other school support	0.5624	0.4882	0.5487
GRPNEI	Participate in neighborhood association	0.5654	0.4818	0.6793
GRPSOC	Participate in social or welfare organization	0.7177	0.6281	0.4393
GRPFRAT	Participate in service or fraternal organization	0.5600	0.4559	0.4947
PUBMEET2	How often attended a public meeting discussing school	0.4968	0.4935	0.4240
VOLTIME2	Number of times volunteered	0.5570	0.5603	0.6493
PROJECT	Worked on community project in past 12 months	0.8170	0.7118	0.3323
BLOOD	Donated blood in past 12 months	0.3781	0.3069	0.5902

Table B2. Factor analysis results for sociability social capital factor.

Variable from SCCBS	Description	Factor Loadings	Correlation between retained factor and original variables	Uniqueness
FRNDHOM2	How often had friends over to your home	0.5091	0.5225	0.4511
FRNDHNG2	How often hung out with friends in a public place	0.4856	0.5033	0.4871
PARADE2	How often attended parade, local sports or arts event	0.4884	0.5025	0.6763
CARDS2	How often played cards or board games with others	0.4075	0.4194	0.6182
FRIENDS	Number of close friends	0.4052	0.3850	0.7667
ARTIST2	How often took part in artistic activity with group	0.4406	0.4521	0.5331
GRPSPORT	Participate in sports club, league, or outdoor activity	0.5323	0.4562	0.3263
GRPVET	Participate in veterans group	0.2894	0.1983	0.4804
GRPELD	Participate in seniors group	0.3527	0.2612	0.4322
GRPART	Participate in literary, art or musical group	0.5455	0.4550	0.3834
GRPHOB	Participate in hobby, investment, or garden club	0.5408	0.4394	0.4810
GRPWWW	Involved in group that meets over the internet	0.3541	0.2241	0.5782
GRPOTHR	Belongs to other kinds of groups or organizations	0.4476	0.3448	0.5925
NEISOC	How often talk with or visit immediate neighbors	0.2595	0.2432	0.6738
CLUBS2	How often attended a club meeting	0.5215	0.5265	0.7125
TEAMSPT2	How often played a team sport	0.4429	0.4586	0.3555

Appendix C: Reduced Form Probit and Wage Determination Estimations.

Table C1. First-stage reduced form probit estimation of being observed in a social occupation.

Variables	Females	Males
High school education=0,1	-0.1005***	0.0529***
	(0.0028)	(0.0055)
Some college education=0,1	0.0765***	0.4406***
	(0.0030)	(0.0056)
College graduate=0,1	0.5912***	0.9227***
	(0.0036)	(0.0063)
Hispanic=0,1	-0.0203***	0.0222***
	(0.0028)	(0.0046)
Black, non-Hispanic=0,1	0.1045***	0.1785***
	(0.0023)	(0.0040)
Other race, non-Hispanic=0,1	-0.1151***	-0.0692***
	(0.0034)	(0.0054)
Live in MSA=0,1	-0.0873***	-0.1176***
	(0.0017)	(0.0027)
Mid Atlantic region=0,1	0.0089***	0.0338***
	(0.0033)	(0.0052)
East North Central region=0,1	-0.0421***	-0.0088*
	(0.0032)	(0.0052)
West North Central region=0,1	-0.0448***	-0.0205***
	(0.0036)	(0.0058)
South Atlantic region=0,1	-0.0452***	-0.0097*
	(0.0033)	(0.0053)
East South Central region=0,1	-0.0794***	0.0031
	(0.0039)	(0.0064)
West South Central region=0,1	-0.0293***	0.0352***
	(0.0035)	(0.0056)
Mountain region=0,1	-0.0551***	0.0116*
	(0.0039)	(0.0061)
Pacific region=0,1	-0.0306***	0.0517***
	(0.0034)	(0.0053)
Age	-0.0192***	-0.0336***
	(0.0004)	(0.0005)
Age squared	0.0002***	0.0004***
	(0.0000)	(0.0000)
Married=0,1	0.0476***	-0.0545***
	(0.0015)	(0.0025)
Citizen=0,1	-0.0146***	0.0250***
	(0.0034)	(0.0052)
disability=0,1	0.0444***	0.0273***
	(0.0020)	(0.0031)
private, not-or-profit employer=0,1	0.1127***	0.6460***

Variables	Females	Males
	(0.0020)	(0.0031)
government employer=0,1	-0.0005	-0.5250***
	(0.0048)	(0.0079)
self-employed=0,1	0.3768***	0.7508***
	(0.0018)	(0.0028)
altruism = medium	-0.0830***	-0.0318***
	(0.0022)	(0.0037)
altruism = high	-0.0570***	-0.0331***
	(0.0031)	(0.0051)
sociability = medium	0.0625***	0.0414***
	(0.0021)	(0.0040)
sociability = high	0.0470***	0.0386***
	(0.0028)	(0.0047)
Number of children in HH[a]	0.0240***	0.0079***
	(0.0006)	(0.0010)
Lived in area 5 yrs or less=0,1[a]	-0.0199***	0.0038
	(0.0016)	(0.0026)
Constant	-1.556***	-2.308***
	(0.0098)	(0.0151)
Observations	9,880, 000	10,440,000

Notes: The dependent variable is an indicator for whether person is observed in a social occupation. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Regression also includes industry dummy variables. Sample sizes are rounded to the nearest 10,000 for disclosure purposes. Among the employed, 22 percent of women and five percent of men are in social occupations.

[a]Indicates regressor is unique to this first-stage estimation (not included in the wage equation).

Table C2. Second-stage OLS wage equation estimations.

	Females		Males	
Variables	Social Occupations	Non-social Occupations	Social Occupations	Non-social Occupations
High school education=0,1	-0.0638***	0.1150***	0.0766***	0.1167***
	(0.0029)	(0.0009)	(0.0071)	(0.0008)
Some college education=0,1	0.2766***	0.2008***	0.5231***	0.1748***
	(0.0027)	(0.0010)	(0.0115)	(0.0009)
College graduate=0,1	0.9724***	0.5326***	1.1198***	0.4974***
	(0.0056)	(0.0014)	(0.0200)	(0.0012)
Hispanic=0,1	-0.0266***	-0.0481***	0.0166***	-0.1076***
	(0.0022)	(0.0010)	(0.0042)	(0.0009)
Black, non-Hispanic=0,1	0.1128***	-0.0123***	0.1357***	-0.1092***
	(0.0020)	(0.0009)	(0.0049)	(0.0009)
Other race, non-Hispanic=0,1	-0.0025	-0.0029**	0.0098*	-0.0419***
	(0.0029)	(0.0013)	(0.0052)	(0.0012)
Live in MSA=0,1	0.0334***	0.1754***	0.0162***	0.1473***
	(0.0015)	(0.0006)	(0.0032)	(0.0006)
Mid Atlantic region=0,1	0.0225***	-0.0168***	0.0499***	-0.0063***
	(0.0022)	(0.0012)	(0.0045)	(0.0012)
East North Central region=0,1	-0.1099***	-0.0918***	-0.0588***	-0.0438***
	(0.0022)	(0.0012)	(0.0044)	(0.0011)
West North Central region=0,1	-0.1844***	-0.1263***	-0.1575***	-0.1096***
	(0.0024)	(0.0013)	(0.0048)	(0.0012)
South Atlantic region=0,1	-0.1542***	-0.1089***	-0.1212***	-0.1029***
	(0.0023)	(0.0012)	(0.0045)	(0.0012)
East South Central region=0,1	-0.2088***	-0.1692***	-0.1074***	-0.1309***
	(0.0028)	(0.0014)	(0.0055)	(0.0014)
West South Central region=0,1	-0.2259***	-0.1623***	-0.1392***	-0.1280***
	(0.0024)	(0.0013)	(0.0048)	(0.0012)
Mountain region=0,1	-0.2048***	-0.1078***	-0.1425***	-0.0887***
	(0.0028)	(0.0014)	(0.0052)	(0.0013)
Pacific region=0,1	-0.0521***	0.0005	0.0132***	0.0084***
	(0.0024)	(0.0013)	(0.0047)	(0.0012)
Age	0.0296***	0.0425***	0.0312***	0.0557***
	(0.0004)	(0.0002)	(0.0009)	(0.0001)
Age squared	-0.0003***	-0.0004***	-0.0002***	-0.0005***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Married=0,1	0.0491***	0.0114***	0.0749***	0.1670***
	(0.0012)	(0.0006)	(0.0025)	(0.0005)
Citizen=0,1	-0.0050*	0.0682***	0.0273***	0.0657***
	(0.0030)	(0.0013)	(0.0050)	(0.0011)
disability=0,1	-0.0018	-0.0398***	-0.0373***	-0.0584***
	(0.0016)	(0.0007)	(0.0031)	(0.0007)

	Females		Males	
Variables	Social Occupations	Non-social Occupations	Social Occupations	Non-social Occupations
private, not-or-profit employer=0,1	0.0584***	0.0198***	0.4788***	-0.0179***
	(0.0015)	(0.0011)	(0.0120)	(0.0017)
government employer=0,1	0.1729***	0.0208***	0.5493***	-0.0129***
	(0.0023)	(0.0011)	(0.0123)	(0.0011)
self-employed=0,1	-0.3224***	0.0173***	-0.6544***	0.0111***
	(0.0061)	(0.0022)	(0.0189)	(0.0014)
altruism = medium	-0.0242***	0.0469***	0.0149***	0.0500***
	(0.0019)	(0.0007)	(0.0036)	(0.0006)
altruism = high	0.0680***	0.0870***	0.0847***	0.1204***
	(0.0022)	(0.0010)	(0.0043)	(0.0010)
sociability = medium	0.1280***	0.0540***	0.1073***	0.0795***
	(0.0016)	(0.0008)	(0.0038)	(0.0008)
sociability = high	0.0908***	0.0170***	0.1270***	0.0595***
	(0.0019)	(0.0011)	(0.0042)	(0.0010)
$\hat{\lambda}_i$	1.2747***	0.0553***	1.1770***	0.2401***
	(0.0151)	(0.0039)	(0.0280)	(0.0042)
Constant	-2.1821***	1.1490***	-2.5029***	1.0315***
	(0.0416)	(0.0037)	(0.0873)	(0.0033)
Observations	2,150,000	7,740,000	560,000	9,880,000
R-squared	0.2459	0.2233	0.2619	0.2861

Notes: Dependent variable is log hourly wage. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Regression also includes industry dummy variables. Sample sizes are rounded to the nearest 10,000 for disclosure purposes.